Estimating NVP variability for deterministic models

van Groenendaal, W.J.H.

Published in:
European Journal of Operational Research

Publication date:
1998

ELSEVIER

European Journal of Operational Research 107 (1998) 202–213

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

Theory and Methodology

Estimating NPV variability for deterministic models

Willem J.H. van Groenendaal *

Faculty of Economics and Business Administration, Tilburg University, P.O. box 90153, 5000 LE Tilburg, The Netherlands

Received 1 March 1996; revised 1 March 1997

Abstract

For decision makers the variability in the net present value (NPV) of an investment project is an indication of the project's risk. Risk analysis is one way to estimate this variability. However, risk analysis requires knowledge about the joint-distribution function of the inputs. Modeling this is often very difficult if not impossible for large long-term investment projects, such as energy infrastructures. In practice, the analysis of the variability is then usually restricted to deterministic sensitivity analysis, such as 'one-factor-at-a-time' and scenario analysis. These deterministic analyses, however, do not account for the total variability in the NPV. This paper shows that the use of experimental design (taken from statistical theory) in combination with regression (meta) modeling is a better approach. © 1998 Elsevier Science B.V.

Keywords: Experimental design; Metamodels; Risk; Sensitivity analysis; Net present value

1. Introduction

For the assessment of the feasibility of large investment projects, such as gas transmission systems, often large non-linear models are built. In general, only one solution for the investment problem is designed (Mintzberg et al., 1976; Van Groenendaal, 1995a,b). An important criterion to evaluate an investment project is its net present value (NPV). However, if for the most likely or base case scenario the project is profitable (NPV > 0), information on the *robustness* of the base case result is needed to support the decision makers in their assessment of a project's risk. To support risk assessment *risk analysis* is often proposed, that is, the determination of the probability distribution of the NPV. To obtain this distribution, information is required on the distribution of the input variables and model parameters (further referred to as factors). For large projects, however, modeling the joint distribution of even a limited number of input variables and parameters can be rather problematic, if not impossible. This is especially true for investments in large energy infrastructures, due to the versatile nature of the energy market. As a result of these problems, the analysis of the robustness is often restricted to *one-factor-at-a-time sensitivity analysis* or the analysis of a limited number of *scenarios*, where the term 'sensitivity analysis' indicates deterministic approaches to the problem. Note that the one-factor-at-a-time approach is the most popular one in applied work; for example, see Duvigneau and Prasad (1984), UNIDO (1986), and Ward and Deren (1991).

Furthermore, both deterministic methods result in limited and often misleading information about the

* Fax: +31-13-4663377; e-mail: groenend@kub.nl

PII S0377-2217(97)00138-0

Table 1
Plackett–Burman design for ten factors

Combination	Factor											
	1	2	3	4	5	6	7	8	9	10	11	12
1 Investment costs	+	+	−	+	−	+	−	−	−	+	−	−
2 Construction time	−	+	+	−	−	+	+	−	−	−	+	−
3 Reserves West Java	+	−	+	+	−	+	+	+	−	−	−	−
4 Real GVA	−	+	−	+	−	−	+	+	+	−	−	−
5 Energy price levels	−	−	+	−	−	+	−	+	+	+	−	−
6 Relative gas/oil price	−	−	−	+	−	+	+	−	+	+	+	−
7 Gas purchase prices	+	−	−	−	−	−	+	+	−	+	+	−
8 Coal prices	+	+	−	−	−	+	−	+	+	−	+	−
9 Other costs	+	+	+	−	−	−	+	−	+	+	−	−
10 Discount rate	−	+	+	+	−	−	−	+	−	+	+	−

Source: Kleijnen (1975, p. 332).

variability of the NPV, especially the one-factor-at-time-approach, which is crude, and is known to be naive and inefficient from a statistical point of view (Kleijnen and Van Groenendaal, 1992, pp. 171–2). More reliable information about the variability of the NPV can be obtained through the *design of experiments* (DoE) taken from statistical theory.

In this paper a three-step procedure is presented to improve traditional sensitivity analysis through the application of experimental design theory in combination with regression metamodeling. It will be shown that the use of experimental designs leads to more adequate information about the variability in the NPV, and that regression metamodeling can be used to evaluate the variability. The procedure developed requires no knowledge about the distribution of the input variables or the parameters of the simulation model, has the advantage that it is easy to apply, and leads to better results than does traditional sensitivity analysis.

A case-study, the investment in gas transmission on the Indonesian island of Java, is used to illustrate the approach. This project requires an overall investment of U.S. $1.9 billion in 1990 prices, and will be realized in three separate investment phases. Phase I is currently executed, and aims at utilizing the gas reserves in the vicinity of Java to develop the market. Phase II will start somewhere between 2005 and 2009. By then, the gas reserves in the vicinity of Java will be fully utilized, and a connection is needed to gas reserves in East Kalimantan, more than 1,000 miles away. Phase III of the investment schedule is the installation of compressor stations to boost the system's capacity. For the base case scenario the NPV for the project is 898.5 billion Indonesian rupiah in 1990 prices (U.S. $\$1 \approx 1770$ rupiahs), so the project is feasible. Ten factors affect the NPV (see the first column of Table 1); they present other projects or government policies (diversifiable risk) as well as indicators for economic development, such as energy prices and economic growth (non-diversifiable risk). The results of the three-step procedure for this project give the Indonesian government clear insight into the variability of the NPV and thus the risks of investing in gas transmission on Java.

This paper is organized as follows. Section 2 contains an outline for a three-step procedure for the determination of NPV variability. Section 3 applies the procedure for investment in gas transmission on the island of Java. Section 4 contains conclusions.

2. Determination of NPV variability using deterministic simulation

The first step in the estimation of the variability in the NPV is the determination of the effect of a positive and a negative factor change. Normally, decision makers are more interested in information on what can go wrong than in information on 'pleasant surprises'. For the analysis of NPV variability this implies a special interest in factor changes that have a negative effect on the NPV. For the base-case scenario the case-study is financially sound: the NPV is 898.5 billion rupiahs. The next thing decision makers want to know is: does the project go wrong when the actual factor value(s) deviate from the base-case values? Hence, deviations from the base-case values that will improve the result are less interesting (a positive NPV is sufficient to accept the project). The decision makers want to know under what circumstances the decision to invest will be incorrect, and they wish to judge the likelihood of these unfavorable circumstances. To determine factor changes that cause negative effects the first step is to apply one-factor-at-a-time sensitivity analysis.

The first step of our three-step procedure is required only when it is not known how factors affect the NPV, which in general will be the case for large investment projects. For example, a larger economic growth will, among other things, lead to more gas sales, but also to an earlier construction of the second and third phases of the project. More sales will have a positive effect on the NPV, but earlier investment in a connection to gas resources further away will affect the NPV negatively. The overall effect of increased economic growth is therefore not known in advance.

An advantage of the one-factor-at-a-time approach is that it is easy to understand, because it treats every factor separately. However, changing *one-factor-at-a-time* is also its main weakness, because it does not allow checking for factor interactions or combined effects. To obtain information on both main effects (as the one-factor-at-a-time approach does) and interactions, other experimental designs are necessary. The full factorial design is a design which allows the estimation of *all* possible interactions (Kleijnen and Van Groenendaal, 1992, pp. 172–4). However, with every factor at two levels only, it requires 2^k simulation runs for k factors; so ten factors require $2^{10} = 1024$ simulation runs. In other words, a full factorial design – even with a moderate number of factors – may require too many simulation runs; it is not a practical solution in case of many factors and long simulation runs. The large number of simulation runs was, however, one of the reasons why risk analysis was proposed (Pouliquen, 1970). Furthermore, since the simulation model is large, every simulation requires substantial computer time. This is not feasible in a commercial setting for project evaluation. Moreover, except for the pathological case where all factors interact, a full factorial design is not necessary. In most cases, accounting for two factor interactions is sufficient; for example, see Kleijnen and Standridge (1988).

There are experimental designs that allow estimating all main effects and a limited number of interactions, but require much less than 2^n simulation runs. So experimental design theory can help to meet the information needs of the decision makers. This alternative has been neglected in project appraisal (none of the many guidelines for project evaluation used by international lending agencies mentions this approach).

What is needed is an approach that allows estimating *all main effects*, and possibly *some interactions*. Suppose for the time being, that there are k factors and no interactions. Then the *full factorial* design needs 2^k observations to estimate only $k+1$ effects, namely k main effects plus the overall mean. In principle $k+1$ observations suffice to estimate $k+1$ effects (Kleijnen and Van Groenendaal, 1992, p. 175). In other words, for the $k+1$ effects it suffices to simulate a fraction, namely 2^{-p}, of the 2^k observations. These 2^{k-p} designs are called fractional factorial designs. Obviously, if only main effects are important the minimum number of simulation runs is obtained by choosing the maximum p such that $2^{k-p} \geq k+1$ still holds. For the case-study $k+1 = 11$, which calls for a 2^{10-6} ($=16$) design.

Fractional factorial designs have a number of runs equal to a power of two. So when the number of factors becomes large, the number of simulation runs is still large. A class of designs that allows a more gradual increase in the number of simulation runs is the *Plackett–Burman* design type. These designs require a number of runs equal to a multiple of four. Hence, for ten factors the Plackett–Burman design with 12 runs can be used (Kleijnen, 1975, pp. 329–36). The transposed Plackett–Burman design matrix (say) X^T is given in Table 1. Every column in that table represents a simulation run. A plus sign (+) stands for the base case value of the corresponding factor, and a minus sign (–) for the value that has a negative influence on the base case result. It is easy to see that the columns of the Plackett–Burman design matrix are orthogonal; thus $(X^T X)^{-1} = 12^{-1} I$, where $I \in R^{10 \times 10}$ denotes the identity matrix. Furthermore, the design satisfies one linear constraint: the sum of the first 11 rows of X equals minus row 12. The augmented matrix $(e : X)$, with $e = (1,1,\dots,1) \in R^{12}$ has the same properties as the matrix X. Note that the rows of the design matrix (or columns in Table 1) can be interpreted as scenarios, some of which may make economic sense, others being less likely.

The last step in our three-step procedure is analyzing the data obtained in the second step by fitting a regression metamodel; that is, approximate the I/O behavior of the (complicated) simulation model through a regression metamodel (Kleijnen and Van Groenendaal, 1992, pp. 149–50). Assuming that only main effects and two factor interactions are important, the metamodel has the form

$$\underline{y}_i = \beta_0 + \sum_{j=1}^{10} \beta_j x_{ij} + \sum_{j=1}^{10} \sum_{k=j+1}^{10} \beta_{jk} x_{ij} x_{ik} + \underline{\varepsilon}_i;$$

$$i = 1,2,\dots,12, \tag{1}$$

where $\underline{y}_i$ denotes the result of the ith simulation run $\underline{\mathrm{NPV}}_i$, x_{ij} denotes the value of factor j in simulation run i, $\underline{\varepsilon}_i$ denotes the approximation error, which is assumed to be 'white noise', that is, $\underline{\varepsilon}_i$ is additive normally independently distributed noise with $\mathrm{E}(\underline{\varepsilon}_i) = 0$ and $\mathrm{V}(\underline{\varepsilon}_i) = \sigma^2$; and the underscore denotes random variables. The 'white noise' assumption allows the use of simple statistical analysis and tests (such as the F- and t-tests). Alternatively, one can drop this assumption, and use the Ordinary Least Squares (OLS) curve fitting, which yields less powerful analysis (such as R^2).

Note that the choice of the regression metamodel and the experimental design are closely related. A relationship which is insufficiently emphasized in the standard literature on DoE (Kleijnen, 1996). The starting point for the design is the wish to estimate the main effects and the expectation that only a limited number of two-factor interactions in (1) are significant. Only in case the results obtained are unsatisfactory, the design has to be augmented.

The choice of the experimental area affects this analysis. From a statistical point of view, one should stick to the total experimental area, that is, use the widest range in factor levels. This will give the best variance estimator for the metamodel. The metamodel, however, is only an approximation of the simulation results, and the fit will be better when the experimental area is smaller. Here the latter argument is more important than the former. So there is also a methodological motivation for the analysis of negative effects only.

If there are no interactions we can estimate the coefficients β_0 and $\beta_M = (\beta_1,\dots,\beta_{10})^{\mathrm{T}}$ in (1) by applying OLS (OLS gives Best Linear Unbiased Estimators if $\underline{\varepsilon}$ is white noise): $\beta_0 = 12^{-1}\sum_{i=1}^{12} y_i$, because $e^{\mathrm{T}}X = \underline{0}$ since every column in X contains an equal number of $+1$ and -1 entries; $\beta_M = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y$. However, if the assumption of no interactions is false, our estimates are biased, as is shown next.

Let $\beta_A = (\beta_{1,2},\dots,\beta_{1,10},\beta_{2,3},\dots,\beta_{9,10})^{\mathrm{T}} \in R^{45}$ be the vector of two factor interactions. Have the matrix of independent variables associated with β_A be denoted by $V = (V_1,V_2,\dots,V_9) \in R^{12\times 45}$ with $V_i = (x_i x_{i+1}, x_i x_{i+2},\dots,x_i x_{10})$. $x_1 x_2$ denotes the 12-dimensional vector with elements $x_{j1}x_{j2}$, $j = 1,2,\dots,12$. Then the expected value of $\underline{\beta}_M$ is

$$\begin{aligned}
\mathrm{E}\left(\underline{\hat{\beta}}_M\right) &= (X^{\mathrm{T}}X)^{-1} X^{\mathrm{T}} (X\beta_M + V\beta_A) \\
&= \beta_M + (X^{\mathrm{T}}X)^{-1} X^{\mathrm{T}} V\beta_A \\
&= \beta_M + 12^{-1} X^{\mathrm{T}} V\beta_A .
\end{aligned} \tag{2}$$

The matrix $12^{-1}X^{\mathrm{T}}V$ is called the *alias* or bias matrix; this matrix shows how the main effects are confounded with the interactions. (For a comprehensive treatment of the subject see Raktoe et al. (1981).)

The estimator for the main effects $\underline{\beta}_M$ would be unbiased if $X^{\mathrm{T}}V = 0$. Unfortunately $X^{\mathrm{T}}V = 0$ does not hold for the Plackett–Burman design; fortunately this equation can be achieved by applying the Box–Wilson *foldover* theorem; see Box and Wilson (1951) and Kleijnen (1975, p. 342–344). A foldover is obtained by adding $-X$ to the original design matrix X, so 24 instead of 12 simulation runs are executed. Through this addition of $-X$ to X the OLS estimator for β_M is no longer biased by interactions between two factors (Kleijnen, 1975, pp. 413–4). A design in which no main effect is confounded with any other main effect or any two-factor interaction, but in which the two-factor interactions are confounded with each other is called resolution IV design (Kleijnen, 1975, p. 329). (Obviously no unbiased estimators of all main effects and two-factor interactions are possible: there are $1 + 10 + (10 \times$

$9/2$) effects and only 24 runs.) Applying the foldover technique changes (1) into

$$y = \left(e_{24}; \begin{pmatrix} X \\ -X \end{pmatrix}; \begin{pmatrix} V \\ V \end{pmatrix}\right) \begin{pmatrix} \beta_0 \\ \beta_M \\ \beta_A \end{pmatrix} + \underline{\varepsilon} = Z\beta + \underline{\varepsilon}. \quad (3)$$

The OLS estimator $\underline{\beta}$ follows from (3): $\underline{\beta} = (Z^TZ)^{-1}Z^T\underline{y}$. The matrix Z^TZ is

$$Z^TZ = \begin{pmatrix} 24 & 0 & 0 \\ 0 & 24I & 0 \\ 0 & 0 & 2V^TV \end{pmatrix}. \quad (4)$$

This is easy to prove; remember that every column in X and V contains as many -1 as $+1$ elements, and that X is an orthogonal matrix.

The maximum number of two-factor interactions that can be estimated with the foldover is equal to the rank of $(V:V)^T$, which is equal to the rank of V. Because V consists of combinations of columns of X, the rank of V cannot exceed the row rank of the full Plackett–Burman design matrix for 12 factors, which is only 11 since the addition of the first 11 rows of this design matrix is equal to minus row 12. If we calculate the rank of V it is actually 11. *This rank implies that we can estimate up to 11 two-factor interactions.*

The structure of Z^TZ allows the computation of β_0 independently of β_M and B_A, namely $\beta_0 = \Sigma_{i=1}^{24} y_i/24$. Similarly, $\beta_M = 24^{-1}X^Ty$. Next, 11 independent columns from V need to be selected to form the matrix V_I. The rest of the columns of V are combined in the matrix (say) V_A. The resulting alias matrix $(V_I^TV_I)^{-1}V_I^TV_A$ for the 11 interactions can be formed in the same way as for $\underline{\beta_M}$ in (2); also see Kleijnen (1975, pp. 415–6).

Note that if one wants estimates of more than 11 two-factor interactions or higher order interactions, the design needs to be augmented (beyond the foldover), and new simulation runs are required. Next, this three-step procedure is applied to estimate the variability in NPV.

3. Applying the three-step procedure for the case study

The results of the first step, one-factor-at-a-time analysis, are given in Table 2, which allows us to determine those deviations that have a negative effect on the NPV. The factor values in Table 2 are chosen in such a way that the experimental area is as large as possible; that is, the borders of the experimental area are the extreme values for the variables considered. By choosing the values associated with factor changes in this way, the magnitudes of the estimated β_j's in (1) are comparable and show the importance of the effects (Bettonvil and Kleijnen, 1990).

Note that all entries in Table 2 are positive, *so no single factor change causes the project to go wrong.*

Table 2
One-factor-at-a-time sensitivity analysis for NPV

Factor j	Value	NPV + [a]	NPV −
1 Investment costs	±25%	659.1	1,138.0
2 Construction time	one year extra	822.0	1,002.9
3 Gas reserves in West Java	±20%	1,141.5	297.8
4 Real GVA	see Appendix 6.1	992.7	793.3
5 Energy price levels	±5%	1,254.1	285.3
6 Relative gas–oil price	±10%	737.9	941.7
7 Gas purchase prices	±5%	472.6	1,324.4
8 Coal price	±3%	956.2	790.1
9 Other costs	±25%	829.9	967.2
10 Discount rate	10% ±2%	577.4	1,425.9
Base-case		898.5	

[a] + and − correspond to a positive and negative change in the input respectively.

For decision makers this information might be misleading when the analysis of NPV variability is restricted to one factor at a time only. Most results are as one would expect: higher (lower) costs lead to lower (higher) NPV values (factors 1, 2, and 9). An increase in the price of coal leads to an increase in the demand for natural gas by the power sector, and thus to a higher NPV.

The change in initial energy prices (factor 5) has the largest effect on NPV. This needs some explanation. Natural gas is priced at fuel oil parity, and the change in initial energy prices implies that relative gas–fuel oil prices remain the same. Therefore the change in volume sold is small for the manufacturing sector. However, given the cost curves for the different technologies of the power sector, the changes in price levels of coal and natural gas cause changes in the technology mix for the power sector. The volume effect of this change is considerable. Factor 5 shows the importance of a correct, that is, economically efficient, pricing of energy.

Changing the price of natural gas relative to the price of oil (factor 6) has only a minor effect on the NPV. Lowering the relative gas–oil price, however, has a positive effect. The cause is that lowering the relative price leads to the conversion of existing production processes to natural gas, but does not affect the demand by new investments. In the base case natural gas is priced at fuel oil parity (154 rupiah/m^3). Pricing gas at fuel oil parity means that the prices are equal with respect to their heat content. Because natural gas requires less investments and is less contaminating nearly all new investments will use gas at this price, whereas due to sunk cost most existing facilities will not convert to gas; see also Van Groenendaal (1997). Fig. 1 shows the gas–oil price sensitivity pattern for existing production processes. If the price is lowered to 152 rupiah, at least some food processing factories will switch to natural gas. At 150 rupiah textile and cardboard production change to gas; the last big increase in gas demand is between 146 and 144 rupiah/m^3, when paper and pulp production switches to natural gas. (The bar at 130 rupiah/m^3 in Fig. 1 indicates the maximum demand for the period displayed.) All changes in demand discussed are within the −10% relative price decrease; the increase in demand volume offsets the decrease in price which explains the increase

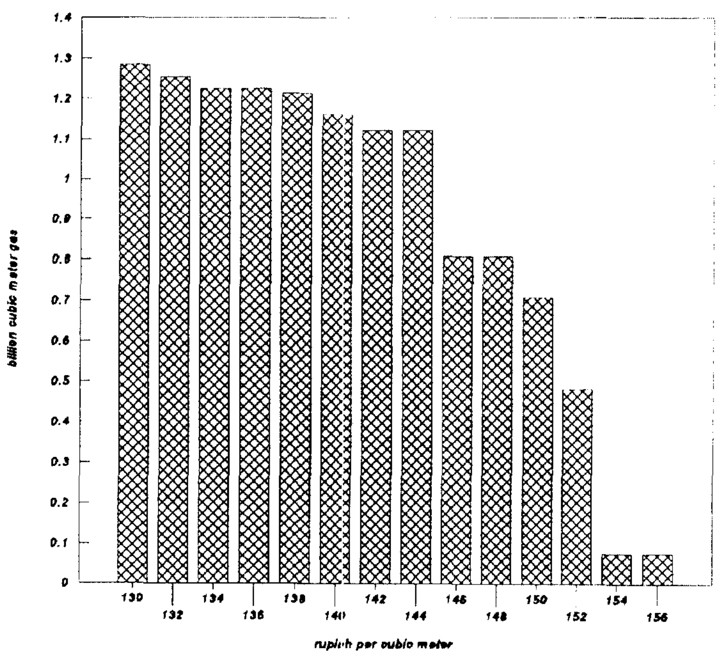

Fig. 1. Demand-price sensitivity of manufacturing sector.

Table 3
Estimation results for a one factor at a time design

	Full experimental area	Negative area only
$\hat{\beta}_0$	871.8	898.5
$\hat{\beta}_1$	119.7	119.7
$\hat{\beta}_2$	45.2	38.3
$\hat{\beta}_3$	210.9	300.4
$\hat{\beta}_4$	49.9	52.6
$\hat{\beta}_5$	242.2	306.6
$\hat{\beta}_6$	51.0	80.3
$\hat{\beta}_7$	213.0	213.0
$\hat{\beta}_8$	41.5	34.3
$\hat{\beta}_9$	34.3	54.2
$\hat{\beta}_{10}$	212.1	160.6
R^2_{adj} [a]	0.98	[b]
DF [c]	10	0

[a] $R^2_{\mathrm{adj}} = [R^2 - (k/(m-1))][(m-1)/(m-k-1)]$, with m the number of observations, and k the number of factors (excluding $x_0 \equiv 1$).
[b] Not defined for $m = k$.
[c] DF: degrees of freedom $m - k - 1$.

in NPV. (Apparently the price elasticity of demand is larger than -1 in this price range.)

Note that analysts often simulate more than two extra levels in one-factor-at-a-time sensitivity analysis. In general, intermediate simulation runs will add no extra information, unless there are very strong non-linearities; see also Eschenbach (1992).

To analyze the effect of the size of the experimental area the data of Table 2 are used to estimate the main effects for the one factor at a time approach, that is, estimate $\beta_M = (\beta_0, \beta_1, \ldots, \beta_{10})^{\mathrm{T}}$. By using the values $+2$ and -2 (instead of the more familiar values $+1$ and -1) for the factor changes for the total experimental area, the estimates are scaled in such a way that they become comparable with the results for the negative area only (on which the next step of our three-step procedure concentrates). Without this scaling the estimates cannot be compared.

The OLS result $\hat{\beta}_M = (\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_{10})^{\mathrm{T}}$ for the 11 coefficients is given in Table 3 under the heading 'full experimental area'. Because $X^{\mathrm{T}}X$ is a diagonal matrix, there are only two values for the standard deviations: $\hat{\sigma}_{\hat{\beta}_0} = 22.0$ and $\hat{\sigma}_{\hat{\beta}_j} = 35.7$ ($j = 1, 2, \ldots, 10$). With $21 - 11 = 10$ degrees of freedom, the estimates for β_0, β_1, β_3, β_5, β_7, and β_{10} in the NPV model are significant at the level $\alpha = 0.01$ ($t_{10;0.01} = 2.76$). β_4 and β_6 are only just significant at the $\alpha = 0.10$ level ($t_{10;0.10} = 1.37$). All other estimates are insignificant.

The one-factor-at-a-time approach uses 21 observations to estimate 11 parameters. However, 11 simulation runs would have been sufficient to estimate the main effects. Because decision makers are more interested in negative effects than in windfalls, the estimate of $\hat{\beta}_M$ based on these data is also given in Table 3 under the heading 'negative area only'.

The latter estimates show a slightly different ranking of factors. For factor 10 (the discount rate) a

Table 4
NPV results for the Plackett–Burman design (X) and its foldover ($-X$)

Combination	Factor 1	2	3	4	5	6	7	8	9	10	NPV X	$-X$
1	+	−	+	−	−	−	+	+	+	−	130.6	−370.8
2	+	+	−	+	−	−	−	+	+	+	−1,017.8	125.4
3	−	+	+	−	+	−	−	−	+	+	−189.4	−554.2
4	+	−	+	+	−	+	−	−	−	+	−271.9	20.5
5	+	+	−	+	+	−	+	−	−	−	191.3	−572.2
6	+	+	+	−	+	+	−	+	−	−	208.0	−672.2
7	−	+	+	+	−	+	+	−	+	−	−154.5	−90.7
8	−	−	+	+	+	−	+	+	−	+	446.9	−842.7
9	−	−	−	+	+	+	−	+	+	−	−609.9	220.2
10	+	−	−	−	+	+	+	−	+	+	294.9	−588.2
11	−	+	−	−	−	+	+	+	−	+	−879.8	179.8
12	−	−	−	−	−	−	−	−	−	−	−730.4	898.5

difference in magnitude is obvious, because of its non-linear effect in the calculation of NPV. The differences between the two estimates for the factors 3 and 5 indicate that the metamodel is inappropriate; that is, higher order effects or interactions between factors might be important.

Remark. The literature on risk analysis has suggested other methods to detect what factors are important. To rank factors Hull (1980, pp. 120–1) introduces the term range coefficient for $(\mathrm{NPV}_{j\mathrm{H}} - \mathrm{NPV}_{j\mathrm{L}})2/[\max_j(\mathrm{NPV}_{j\mathrm{H}} - \mathrm{NPV}_{j\mathrm{L}})^2]$, with $\mathrm{NPV}_{j\mathrm{H}}$ and $\mathrm{NPV}_{j\mathrm{L}}$ the maximum H and minimum L value for factor j respectively. It is easy to see that $(\mathrm{NPV}_{j\mathrm{H}} - \mathrm{NPV}_{j\mathrm{L}}) = 4\hat{\beta}_j$ ($\hat{\beta}_j$ the OLS estimate of the jth main effect). Because taking the square of a non-negative variable is a monotone transformation, the range coefficient will lead to the same ranking as does our approach. Applying Hull's criterion to eliminate factors at a level of 0.10 leads to the same result: the factors 1, 3, 5, 7, and 10 are important. So both methods identify the same significant factors. The regression approach is, however, better known and understood.

The results of the 24 simulation runs for NPV using the foldover of the Plackett–Burman design are shown in Table 4. Combinations 1 to 12 are identical to the 12 columns of Table 1. Remember that our analysis focuses on those conditions that will jeopardize the positive advice for the investment project that follows from the positive NPV value for the base case. Hence, all results are expected to be worse than the base-case result. This explains the many (14) negative entries in the columns denoted by X and $-X$ of Table 4. Each single minus value of the ten factors lowers the NPV; so only if the interactions of factors are positive, are these negative effects mitigated. When the interaction of factors strengthens the main effects, the results of the simulation will be even more negative.

For example, in combination 12 all factors are at their 'minus' level; hence the NPV is negative (namely, −730.4). However, all factors at their 'minus' level is not the largest negative result. The effect of a larger discount rate (factor 10) can mitigate the effect of negative cash flows. This occurs for combinations 2 and 11 in the column denoted by X, and for combination 8 in the column denoted by $-X$.

Remark. Many companies use the 'worts-case' scenario as a benchmark. In case the worst-case scenario is interpreted as $(-1, -1, \ldots, -1)$, the results in Table 4 show that the presence of interactions might lead to the wrong conclusion.

3.1. Estimation of main effects and important two-factor interactions

The analysis starts with the OLS estimation of the main effects for both the original Plackett–Burman design and the foldover, using the NPV data of Table 4. The estimation results in Table 5, show that the foldover does not have much effect on the estimated main effects. The main effects that are important when using the data of the Plackett–Burman design only, are in order of importance β_5, β_3, β_7, and β_1 (besides the overall mean β_0). When the foldover is used, these effects remain the most important ones; the t-values for these five coefficients are all significant at the level $\alpha = 0.01$ ($t^{0.01}_{13=24-11} = 2.85$). A comparison of Tables 3 and 5 shows that, with the exception of $\hat{\beta}_{10}$, the same main factors are important.

Deleting those regressors whose estimated main effects do not significantly differ from zero at $\alpha = 0.05$, increases the R^2_{adj} from 0.87 to 0.91 (not displayed in the table). Note that because $X^{\mathrm{T}}X = 24I$ is a diagonal matrix, the values of the estimates of the main effects do not change when regressors are deleted or added.

Once it is known which main effects matter, it is possible to determine whether or not there are significant two-factor interactions by estimating a model with only the significant main effects plus the first 11 independent interactions: $(\hat{\beta}_{1,2}, \hat{\beta}_{1,3}, \ldots, \hat{\beta}_{1,10}, \hat{\beta}_{2,3}, \hat{\beta}_{2,6}) \in R^{11}$. This raises the R^2_{adj} from 0.90 to 0.97. The estimates for the interactions and R^2_{adj} are also displayed in Table 5 (again, because X is an orthogonal matrix the estimates for the main effects remain the same). The increase in R^2_{adj} suggest that there are significant interactions; several interactions have the same magnitudes as some main effects have.

A more formal test on interactions is the F-test for the hypothesis H_0: $\beta_{1,2} = \beta_{1,3} = \cdots = \beta_{2,6} = 0$; see Kleijnen (1975, pp. 727–30). The value of the F-statistic is 4.80, which is significant at the 5% level, since $F^{11}_{8;0.05} = 3.31$ (the degrees of freedom (DF) are explained as follows: 11 interactions are hypothesized to be 0, and $24 - (5 + 11) = 8$ degrees of freedom are left for the full model). So the H_0 hypothesis is rejected for the NPV model, pointing at interaction effects.

Note that even if H_0 is accepted, this only means zero *weighted sums* of two-factor interaction effects that can be obtained from the alias matrix; see Van Groenendaal (1997, p. 223). *Failure to reject H_0 does not prove that there are no interactions*!

There are $\sum_{j=1}^{11} \binom{45}{j}$ possible models with two-factor interactions. So determining which individual interactions are significant is problematic without making further assumptions. An assumption often made in Risk Analysis is to limit the interactions to those between factors with *significant* main effects. For the NPV model this refers to the interactions between investment costs and West Java's reserves ($\beta_{1,3}$), investment costs and all energy prices ($\beta_{1,5}$), investment costs and the purchase prices ($\beta_{1,7}$), West Java's reserves and down-stream energy price levels ($\beta_{3,5}$), West Java's reserves and gas purchase prices ($\beta_{3,7}$), and energy price levels in combination with a change in gas purchase prices ($\beta_{5,7}$). An estimation of this model, and testing it against the model without interactions shows that there are indeed some significant interactions. However, the F-statistic is 3.14, which is barely significant at the 5% level ($F^{6}_{13;0.05} = 2.92$). This result is not very satisfactory.

Taking into account those interactions that are related to significant main effects is a reasonable approach when no other information is available; the simulation model is then treated as a black box. In this case (and most other applications), however, this approach is not necessary. Clues can be derived from the simulation model itself, and from the intermediate simulation results that lead to the NPV values in Table 4. That table gives only the NPV, not any more details on the simulation runs. In the simulation model, economic growth plays a role in nearly every submodel. However, Table 5 shows that economic growth (factor 4) has no significant effect on the NPV. This seems odd, and is also contrary to what economic theory tells us. Studying the detailed simulation results of the 24 simulation runs shows that economic growth does strengthen the effect of a change in the West Java reserves (factor 3), but it also strengthens the effects of some of the changes in prices (factors 5, 6, 7, and 8). Therefore the search for interactions is augmented to these six variables (factors 3–8), so there are fifteen potential two-factor interactions. After some analyses a combination

Table 5
Estimated overall mean, main effects, and two-factor interactions

	Plackett–Burman (X)	Foldover (X: $-X$)		Two factor interactions
$\hat\beta_0$	−215.2	−201.2	$\hat\beta_{12}$	4.9
$\hat\beta_1$	137.7	146.7	$\hat\beta_{13}$	−4.3
$\hat\beta_2$	−91.9	−7.5	$\hat\beta_{14}$	−23.2
$\hat\beta_3$	243.5	237.3	$\hat\beta_{15}$	32.6
$\hat\beta_4$	−20.8	−9.1	$\hat\beta_{16}$	76.2
$\hat\beta_5$	272.1	293.2	$\hat\beta_{17}$	133.6
$\hat\beta_6$	−20.4	−26.8	$\hat\beta_{18}$	−9.3
$\hat\beta_7$	220.1	206.8	$\hat\beta_{19}$	−84.0
$\hat\beta_8$	−71.9	−16.2	$\hat\beta_{1\,10}$	131.6
$\hat\beta_9$	−42.5	−10.0	$\hat\beta_{23}$	−120.1
$\hat\beta_{10}$	−54.4	17.5	$\hat\beta_{26}$	147.2
R^2_{adj}	0.87	0.87		0.97
DF	1	13		8

of the four main effects (see above) and four two-factor interactions give the best result:

$$\hat{NPV} = -201.2 + 146.7x_1 + 237.3x_3 + 293.2x_5 + 206.8x_7 + 70.8x_3x_4 + 102.8x_3x_8 + 106.3x_4x_6 - 38.1x_7x_8. \qquad (5)$$

Its R^2_{adj} is 0.98. The hypothesis H_0: $\beta_{3,4} = \beta_{3,8} = \beta_{4,6} = \beta_{7,8} = 0$ yields $F^4_{15} = 15.09$ and is significant even at the 0.5% level ($F^4_{15;0.005} = 5.80$). Eq. (5) shows that economic growth (x_4) plays a role in the determination of the NPV in combination with West Java's reserves (x_3), which seems logical since faster growth means earlier depletion. Economic growth also affects the NPV in combination with the relative gas–oil price ($\beta_{4,6}$). Faster economic growth leads to more investments in manufacturing, and thus to a larger demand for gas. Because the application of gas is profitable in these new investments when compared to fuel oil, the willingness to pay in most manufacturing sectors is much higher than the price of gas. As a result the interaction is positive. The significant interaction between the depletion of West Java's reserves and the price of coal ($\beta_{3,8}$) is caused by the increased demand for gas by the power sector when the price of coal increases. The interaction between the purchasing price of natural gas and the price of coal ($\beta_{7,8}$) is less easy to explain. The fact that the constant $\hat{\beta}_0$ differs from zero means that there are some unexplained non-linearities in the simulation results. However, the magnitude of $\hat{\beta}_0$ is moderate and requires no further research. Because the range for every factor is as wide as possible, the magnitude of the coefficients indicates their relative importance.

Note that for the base case (all $x_i = 1$) the value for $\hat{NPV} = 924.6$, which is a good forecast for the base case simulation result: 898.5. Factors mostly are not at their extreme value either; then fractions of -1 or $+1$ can be introduced in (5) to predict the result.

Eq. (5) contains the information on the variability in the NPV and on the most important factors that can cause this variation. This is the kind of information decision makers require. They can learn the following from Eq. (5):

- West Java gas reserves are of crucial importance to the project. These reserves are needed for the development of the market for natural gas in West Java; smaller reserves strongly affect the project's feasibility. Two of the two-factor interactions are related to West Java's reserves. These interactions imply that investment in exploration and exploitation in West Java will be profitable.
- The project is sensitive to energy prices. A correct pricing of energy products is therefore a prerequisite for the success of the project. Most responsive to relative price changes is the power sector. Changing the relative coal–gas price affects the economic viability of the project, mainly through the power sector's fuel-mix. The importance of the power sector is supported by the analysis of an overall change in energy prices, which strongly affects the NPV. In this case too, the effect on the NPV (through the amount of gas sold) results from the fuel-mix for the power sector.
- The demand for natural gas by the manufacturing sector is quite robust with respect to relative gas–oil price changes ($\hat{\beta}_6 = 0$). The reason is that the demand by the non-bulk consumers are high valued applications and make up more than half of the total demand.
- The NPV is strongly affected by purchase price increases (which is equivalent to a reduction in the transmission margin).
- Changes in investment costs have only a moderate influence on the NPV.
- It is also remarkable that economic growth is important only in combination with West Java's reserves and the relative gas–oil price.

3.2. Sensitivity analysis of the variability

The validity of approximation (5) (to the simulation I/O) can be tested through *cross-validation*. Cross-validation means: eliminate factor input combinations x_i (the columns of X and $-X$) and the corresponding simulation result y_i one by one, and re-estimate the regression model (5). The model thus obtained is used to predict y_i for the combination eliminated; this prediction is denoted by $\hat{y}_{-i}$. If the regression metamodel (5) is robust, the prediction $\hat{y}_{-i}$ will be close to y_i. Since there are 24 data points

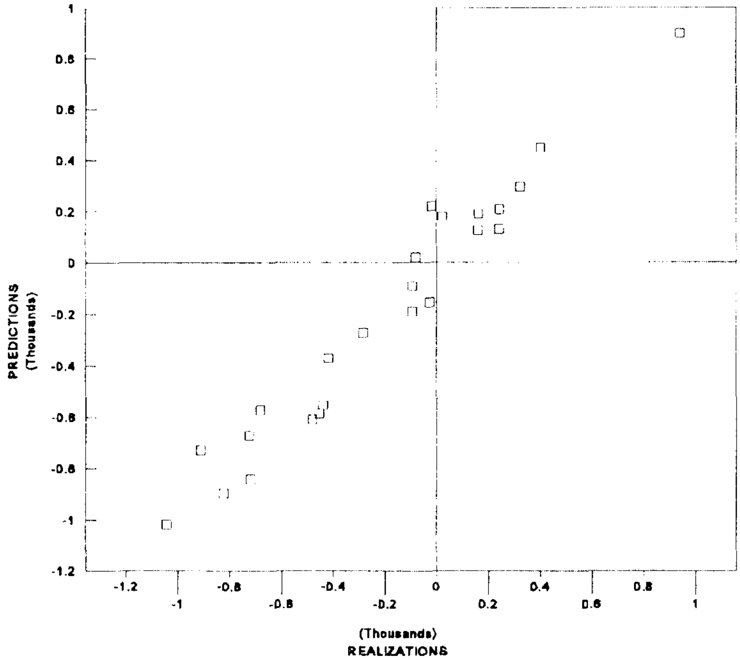

Fig. 2. Scatter plot of NPV regression predictions and simulation realizations

($i = 1,2,\ldots,24$), there are also 24 predictions $\hat{y}_{-i}$. (For more details on cross-validation see Kleijnen and Van Groenendaal, 1992, pp. 156–7.)

To test the robustness for deterministic models Kleijnen and Van Groenendaal (1992, p. 157) suggest computing the relative prediction error $\hat{y}_{-i}/y_i$. However, y denotes NPV, which can change from positive to negative. This switching of sign implies that the relative prediction error is not a good measure for robustness. For example, assume $y_1 = 1$ is predicted by $\hat{y}_{-1} = -10$, and $y_2 = 1{,}000$ by $\hat{y}_{-2} = 1{,}500$. The first relative prediction error is -10, and the second 1.5, so the first error would be regarded as more severe than the second one. The absolute difference $|\hat{y}_{-i} - y_i|$ for the first prediction, however, is only 11 compared to 500 for the second one. It is obvious that in this case the second prediction error is worse than the first. To indicate the quality of the predictions $\hat{y}_{-i}$ obtained through cross-validation, a scatter plot is more appropriate here. If the model is perfect, all points in the scatter plot are on a straight line with an angle of 45°. This performance can be quantified by the correlation coefficient between $\hat{y}_{-i}$ and y_i, which is 0.987. Fig. 2 shows that the regression model (5) performs satisfactorily, so no further analysis is required.

4. Conclusion

For many long-term investment projects no information on the probability distribution of the input variables and model parameters is available. As a result risk analysis (or any other probabilistic approach) is not possible. However, decision makers need information on the variability of the project's NPV for the assessment of its risk. Traditionally one-factor-at-a time sensitivity analysis is used, or a limited number of scenarios are analyzed. These two deterministic approaches, however, are insufficient.

To obtain information on the variability of a project's NPV a three-step procedure was developed, based on experimental design theory and regression metamodeling. Experimental design theory in combination with a regression metamodel can be used to design simulation experiments that, through systematic factor variation, result in data on NPV variability. The regression metamodeling was used to gain insight into the information contained in the simulated data. Furthermore, using experimental design theory allows the estimation of interactions between factors, which is not possible in the traditional approaches (nor in risk analysis).

One could argue that the interpretation of the interactions is not supported by a complete statistical analysis, since not *all* factor interactions were systematically checked. However, because the simulation model is not a black box, analysts who understand their problem will normally be able to qualitatively infer where interactions can be expected. The regression metamodel obtained for the variability of the NPV is statistically sound, supported by the knowledge of the problem at hand, and suggested no flaws in our analysis.

The information on the variability in the NPV obtained tells decision makers which factors are important. By carefully choosing the range within which factor values can vary, the magnitude of an estimate $\hat{\beta}_j$ is also an indication of its importance. The experimental area can be restricted to those factor changes that have a negative effect on the NPV, because decision makers are more interested in what can cause the project to go wrong than in pleasant surprises. The smaller experimental area also improves the fit of the metamodel.

Cross-validation was used to test the robustness of the metamodel. This analysis showed that the parameter estimates are stable, so the metamodel is a good one according to this criterion as well.

The presence of two factors which are important only in combination with other factors (interactions), shows that the widely used practice in risk analysis of deleting those factors from the analysis whose main effect is insignificant, can easily lead to errors.

Acknowledgements

The author is grateful to Prof. J.P.C. Kleijnen for his helpful comments on earlier versions of this paper. The author bears sole responsibility for any remaining errors.

References

Bettonvil, B., Kleijnen, J.P.C., 1990. Measurement scales and resolution IV designs: A note. American Journal of Mathematical and Management Sciences 10 (3/4), 309–322.

Box, G.E.P., Wilson, K.B., 1951. On the experimental attainment of optimum conditions. Journal of the Royal Statistical Society, Series B 13 (1), 1–38.

Duvigneau, J.C., Prasad, R.N., 1984. Guidelines for calculating financial and economic rates of return for DFC projects. World Bank Technical Paper Nr. 33, World Bank, Washington, D.C., U.S.A.

Eschenbach, T.G., 1992. Spiderplots versus tornado diagrams for sensitivity analysis. Interfaces 22 (6), 40–46.

Hull, J.C., 1980. The Evaluation of Risk in Business Investment. Pergamon Press, Oxford, Great Britain.

Kleijnen, J.P.C., 1975. Statistical Techniques in Simulation, Part II. Marcel Dekker, New York, U.S.A.

Kleijnen, J.P.C., 1996. Sensitivity analysis and related analyses: A survey of statistical techniques. Journal of Statistical Computation and Simulation (forthcoming).

Kleijnen, J.P.C., Standridge, C.R., 1988. Experimental design and regression analysis in simulation: An FMS case study. European Journal of Operational Research 33, 257–261.

Kleijnen, J.P.C., Van Groenendaal, W.J.H., 1992. Simulation: A Statistical Perspective. John Wiley and Sons, Chichester, England.

Mintzberg, H., Raisinghani, D., Theoret, A., 1976. The structure of the unstructured decision processes. Administrative Science Quarterly 21, 246–275.

Pouliquen, L.Y., 1970. Risk analysis in project appraisal. World Bank Staff Occasional Paper No. 11, International Bank for Reconstruction and Development. John Hopkins Press, Baltimore, U.S.A.

Raktoe, B.L., Hedayat, A., Federer W.T., 1981. Factorial Designs. John Wiley and Sons, New York, U.S.A.

UNIDO, 1986. Manual for the preparation of industrial feasibility studies. United Nations Industrial Development Organization, Vienna, Austria.

Van Groenendaal, W.J.H., 1995a. Supporting strategic decision processes. In: Vlačić, L., Nguyen, T., Ćećez-Kecmanović, D. (Eds.), IFAC/IFIP/IFORS/SEDC Symposium on Modelling and Control of National and Regional Economies 1995. Pergamon Press, pp. 443-50.

Van Groenendaal, W.J.H., 1995b. Assessing demand when introducing a new fuel. Energy Economics 17 (2), 147–161.

Van Groenendaal, W.J.H., 1997. Investing in Energy Market Development. Oxford University Press Energy Series, Oxford University Press, England (forthcoming).

Ward, W.A., Deren, B.J., 1991. The economics of project analysis: A practitioner's guide. The World Bank, Washington, D.C.